a folha

Boletim da língua portuguesa nas instituições europeias

https://ec.europa.eu/translation/portuguese/magazine/pt_magazine_pt.htm

N.º 63 — verão de 2020

Tratar um português

Jorge Madeira Mendes
Antigo funcionário da Direção-Geral da Tradução — Comissão Europeia

Do aristocrático «Vossa Mercê», derivou «você» (depois de uma escala por «vossemecê»).

Entretanto, as voltas que o «você» foi dando tornaram-no plebeu. Passou a forma muito dúbia de a alguém nos dirigirmos (convém sublinhar que me refiro ao falar de Portugal).

Se o interlocutor for indireto, desconhecido (tipo «cliente num discurso publicitário»), ninguém se ofende: «Use o sabonete Bolhinhas — haverá algo de novo em *Você*».

Também pode ser *snob*, tratamento de bom tom no seio das famílias (que se creem) de fina linhagem, quer entre pais e filhos (e vice-versa) quer entre os próprios cônjuges.

Há ainda quem o use como arma de agressão: quando tratamos alguém, corriqueiramente, por «tu», pode ser visto como hostil tratarmo-lo, de repente, por «você». Sei de um chefe de família cujos ataques de acidez o levavam a mimosear com essa forma de veneno a mulher e o filho mais novo.

Por fim, tem uma função de distanciamento: nem o excesso de familiaridade do «tu» nem a deferência injustificável do «Senhor» ou da «Senhora». E aqui as coisas, às vezes, complicam-se: o destinatário de um tratamento por «você» pode ver nele um agente de desprestígio; há o risco de que, indignado, responda «*Você* é estrebaria!» E a única forma de tratarmos alguém por «você» sem o eventual risco de lhe enxovalharmos os pergaminhos parece ser omitindo o deselegante palavrão, mas construindo a frase como se ele lá estivesse. Exemplifico: uma frase como «*Você* devia saber que a sua cara ofende o meu sentido estético» constituirá, quase inevitavelmente, um convite para a refrega; em contrapartida, libertada do «você», tem mais hipóteses de se ficar por uma mera resposta de cariz idêntico: «Devia saber que a sua cara ofende o meu sentido estético».

Se a alguém, em vez de «Têm-me falado de *você*», dissermos «Têm-me falado de *si*», o tom ameniza-se quase sempre. No que toca ao «você», importa, pois, lidar com pinças.

Entre as formas de tratamento intermédias, ou seja, sem o excesso de familiaridade do «tu» nem a deferência do «Senhor», da «Senhora» ou da «Vossa Excelência», a variante menos suscetível de melindrar o interpelado (ou a interpelada) é nomeá-lo(a) como um «ele» (ou uma «ela») ausente: «O que vai a Manuela tomar?», «A que horas quer a mãe deitar-se?», «O Senhor Doutor tem mais consultas hoje?», «Ontem vi o João a passear o seu cão» (note-se que estamos a falar com — e não de — a Manuela, a mãe, o médico e o João, respetivamente).

Humor à parte, a verdade é que não é fácil explicar isto a um estrangeiro, ainda que lusófono de origem (mas não lusitano).

Isto porque o português (de Portugal) é uma língua complicadíssima no que toca a formas de tratamento — possivelmente, um reflexo da nossa falta de descontração (ou da nossa crispação) em matéria de relações humanas.

Na sua maior parte, as línguas das quais conheço alguma coisa são muito mais simples: o espanhol tem o «*tu*» (familiar) e o «*Usted*» (formal e cada vez mais restrito a determinadas esferas), o húngaro tem o «*ti*» (familiar) e o «*Ö*» (igualmente formal e cada vez mais restrito a determinadas esferas), o italiano tem o «*tu*» e o «*Lei*», o francês tem o «*tu*» (ou «*toi*», conforme a posição sintática) e o «*Vous*», o alemão tem o «*Du*» e o «*Sie*», o russo tem o «ты» e o «Вы», o holandês tem o «*je*» (ou «*jij*») e o «*U*» (este último, formal e também cada vez mais restrito a determinadas esferas). O inglês, muito avisadamente, simplificou, nivelando tudo e todos (e até por cima) com o «*you*», que era a forma mais formal (a outra era o familiar «*thou*», que hoje só se encontra em linguagem arcaica e de caráter religioso). Nós temos o «tu», o «você» explícito, o «você» implícito, «o senhor», «a senhora» («o menino» e «a menina»), o «Vossa Excelência», o «Vossa Senhoria», a menção do nome próprio e/ou do título... nalguma coisa havíamos de ser mais ricos.

Ideal é a forma de tratamento dos islandeses: tuteamento com qualquer interlocutor, sendo que o plural (correspondente aos nossos «vós» ou «vocês») só se emprega com mais do que um interlocutor.

Com o hebraico e, consta-me, com o árabe, passa-se algo parecido: qualquer interlocutor é tratado pelos respetivos equivalentes ao nosso «tu» (sem prejuízo de formas de tratamento especiais que se restringem a raríssimos casos de respeito por hierarquias). E os plurais desses equivalentes ao nosso «tu» serão os equivalentes aos nossos «vós» ou «vocês». Repare-se que, nas edições clássicas da Bíblia, há apenas duas formas de tratamento para a segunda pessoa: o singular («tu») e o plural («vós»), devido, muito provavelmente, ao facto de esses textos provirem do latim, do grego clássico e do hebraico (ou aramaico, a variante falada na Palestina ao tempo do Velho e do Novo Testamentos). Entre Deus e Adão, Deus e Noé, Deus e Abraão, Deus e Moisés, Deus e Job, Deus-Pai e Deus-Filho, era tu-cá-tu-lá. E quando, segundo a mitologia cristã, Jesus se despede dos discípulos, diz-lhes: «Não temais, que eu estarei convosco até à consumação dos tempos» (o que, aliás, é bem mais poético do que as calamitosas legendas dos filmes de hoje: «Não tenham medo, que eu irei estar com vocês até ao tempo se acabar»).

Por outro lado, confesso que (quase) me incomoda ser tratado por «tu» por brasileiros empregados em estabelecimentos comerciais, aqui em Portugal: «Queres gelo na bebida?», «Pagas com cartão?». Isto seria inadmissível entre portugueses (a menos que o cliente fosse uma criança). Mas tento relativizar estas situações, porque o mais certo é eles julgarem que estão a ser formais e polidos comigo…

jorge.mendes909@gmail.com

Onde se prova, mais uma vez, que nem tudo o que luz é oiro…

João Miranda
Direção-Geral da Tradução — Comissão Europeia

Embora, no português moderno, se considere que os ditongos *ou* e *oi* são intermutáveis (ouro/oiro, louro/loiro, Douro/Doiro...), nem sempre foi assim. O seguinte trecho da *Crónica Geral de Espanha* mostra que, no séc. XIV, o nome do famoso rio mantinha (pelo menos, para a elite culta...) o *i* herdado do latim *Durius*[1]:

> E entrou el rey dõ Denis ẽ Castella ataa que chegou a Val d'Estilhas e ataa ponte de Doyro, fazendo tudo o que queria.

Mais adiante, ao narrar um episódio da guerra civil que opôs o rei ao seu herdeiro:

> E, quando o iffante vyo como el rey viinha a elle, acolheose a hũu cabeço e o cõcelho de Lixboa e o de Sanctarem começarõ de fugir pera a mata de Loyres.

Pelos vistos, o autor também tinha a noção de que haveria algo[2] em comum entre a várzea próxima de Lisboa e o célebre rio francês (como já referimos num texto anterior[3]...).

Humm... Quase apetece dizer que anda moura (ou moira?) na costa...

Porque, para o cronista, o vil metal é sempre ouro. E isto é perfeitamente normal, uma vez que deriva do latim *aureus*. Por arrasto, o mesmo se aplica as louros cabelos, cuja cor deriva da planta dourada das coroas (e dos temperos), *laureus*[4]. Como não há dous (ou dois?) sem três, nos restaurantes, todos pedem — e bem — dourada, embora alguns achem que esta possa ter reflexos doirados.

Mas passemos a outro registo. Embora em algumas regiões portuguesas, o povo insista, por vezes, em assistir a toiradas, o certo é que, para os romanos, o corajoso e maltratado animal (*taurus*) não tinha nem a pinta de um *i*. Nisto, os minhotos e os galegos andarão mais perto da verdade (e da pronúnica original...), mas «espalham-se» no que toca ao rio. Já agora, na época das máquinas de lavar louça, ainda há muita gente a lavar a loiça, atividade que os romanos fiziam sem *i*, mas com uma sonoridade parecida (*lutea*). Bom, é melhor não nos alargarmos muito com estas cousas (do latim *causa*, que ganhou um *i* para ficar definitivamente coisa a partir do séc. XIX), porque senão damos em doudos, perdão, doidos.

Pois, por muito que a evolução popular da língua esconda o gato etimológico, ele, por vezes, teima em deixar o rabo de fora...

Joao-Manuel.Miranda@ec.europa.eu

---

(1) Opta-se aqui não pelo acusativo, mas pelo nominativo.
(2) D. Pedro de Barcelos não utilizaria aqui esta palavra, que tinha então um significado diferente: «E deu aa reỹa muy grande algo e muitas doas. E deu muy grande algo a todos seus ricos homẽs».
(3) Miranda, J., «Onde se trata de lavradores, figueiras, pinheiros e outras curiosidades pseudoagrícolas», «a folha», n.º 61 — outono de 2019, https://ec.europa.eu/translation/portuguese/magazine/documents/folha61_pt.pdf.
(4) Castanho-dourado, já que, para os romanos, aquilo que conhecemos por louro-claro (invulgar entre os povos mediterrânicos, mas comum entre os povos germânicos, frequentemente escravizados — o mundo dá muitas voltas...) era designado por *fulvus*.

Toponímia da Índia — breve análise

Paulo Correia
Direção-Geral da Tradução — Comissão Europeia



Mombaim he hum rio o mais larguo e de mais fundo que nenhum que aja neste Estado que seja de Sua Magestade. Está de **Baçaim** oito legoas pera o sul e de **Chaul** quatro pera o norte. O rio he de aguoa salgada, pello qual dezembocão e vem ao mar muitos riachos e esteiros do mar que vem daquelles contornos. Não tem nenhum banco, nem restinga ou baixo que a atravece, senão so hũa restinga, que sae da ponta da terra da banda do norte e lança meya legoa ao mar, todo debaixo de pedra que, posto que quando começa logo da terra se deixa ver algũa couza, contudo logo se encobre debaixo d'aguoa, e assy corre emcuberta a dita meya legoa, em muy pouco fundo, de sorte que qualquer embarcação que lhe não dá resguardo toca logo nella. He este rio de Mombaim na entrada de largura de duas legoas e pera dentro logo vay estreitando, mas não muito. Quem vier de fora demandar a barra ha de hir com a proa ao nordeste, rezalvando pella banda do mar o Ilheo que chamão Candil e, por fundo de oito braças, vay pello meio do canal.[(1)]

O contacto direto da língua portuguesa com a Índia e as suas realidades começou em 1498, tendo os cinco séculos dessa viagem sido comemorados com a Expo98. *Os Lusíadas* ou a *Peregrinação* registaram muita toponímia, sobretudo dos territórios indianos então sob domínio português. Os nomes utilizados por Camões, Fernão Mendes Pinto ou outros viajantes da época basearam-se largamente nos endónimos locais e na sua adaptação à língua portuguesa, contrariamente a muitos textos atuais, em que se decalca a grafia do inglês indiano.

As praças-fortes e territórios portugueses do antigo **Estado da Índia** situavam-se maioritariamente na costa ocidental da península Hindustânica em regiões do **golfo de Cambaia** (Damão e Diu) — rodeadas pelo atual estado de Guzerate — e da **costa do Concão** (Baçaim, Salsete[(2)], Bombaim, Chaul, Goa e Mangalor) — nos atuais estados de Maarastra, Goa e Carnataca. A presença portuguesa na maioria dessas posições manteve-se até meados do século XVIII, altura em que ficou reduzida a Goa, Damão e Diu.

Na **costa do Malabar** — no atual estado de Querala — e para lá do cabo Comorim, na **costa do Coromandel** — nos atuais estados de Tâmil Nadu e Andra Pradexe — e no **golfo de Bengala**[(3)] a presença foi mais curta, terminando até meados do século XVII, em ligação com a perda do Ceilão para os holandeses, apenas se mantendo portuguesa na costa leste São Tomé de Meliapor, até meados do século XVIII.

Assim, o contacto foi mais longo com o **concanim**, o **marata** e o **guzerate**, todas línguas da família indo-europeia, sub-ramo indo-ariano, e foi mais efémero com o **canarim**[(4)], o **tâmil** e sobretudo o **malaialim**, línguas dravídicas. Todas estas seis línguas são hoje línguas oficiais ao nível estadual.

Os topónimos portugueses dessas regiões têm uma distinta **sonoridade indiana**, ligada a formas originais encontradas pelos primeiros portugueses chegados à Índia de verter para português os endónimos locais. Não é de excluir a influência que possam ter tido exónimos anteriores a 1498, por exemplo, de árabes ou persas. O caso mais notável é mesmo o exónimo **Índia**, com origem no grego, quando na maioria das línguas indo-arianas o nome é **Bárata** (भारत).

O devanágari

343. (1) The official language of the Union shall be Hindi in Devanagari script.

Constituição da Índia[5]

Neste artigo começarão por analisar-se os endónimos indianos e os topónimos portugueses de territórios onde se verificou presença portuguesa — o designado Estado da Índia —, procurando identificar-se **padrões nas grafias** portuguesas. Usando-se esses mesmos padrões, analisar-se-ão, por fim, os possíveis aportuguesamentos de outros topónimos da geografia indiana (estados, territórios e capitais).

Para tanto, ter-se-á em conta a escrita em **devanágari** dos topónimos e as respetivas pronúncias em hindi[6] — língua oficial da Índia — e noutras língua indianas. O hindi, tal como o concanim e o marata, usa como sistema de escrita esse alfassilabário (ou abugida). Em tabela no final do artigo apresenta-se uma chave simplificada de leitura do devanágari, recorrendo ao alfabeto fonético internacional (AFI), que permitirá analisar mais facilmente os endónimos indianos. O devanágari, além de ter consoantes e vogais independentes, tem vogais inerentes (/ə/) e ligaturas, nomeadamente as ligaturas de consoantes com vogais diacríticas (representando sílabas).

O inglês indiano, cujo papel transitório está consagrado na Constituição da Índia, usa frequentemente o AITS — o alfabeto internacional de transliteração do sânscrito[7], embora omitindo os diacríticos. O AITS é igualmente apresentado no final deste artigo.

Topónimos do antigo Estado da Índia

pt	**hi** **(en)** **en-IN**	**presença portuguesa**	**estado ou território**	**língua**	**gentílico**
Diu	दीव Dio	1535-1961	Dadrá e Nagar Aveli e Damão e Diu	guzerate	diuense
Damão	दमन Daman	1559-1961	Dadrá e Nagar Aveli e Damão e Diu	guzerate	damanense
Dadrá	दादरा Dadra	1779-1954	Dadrá e Nagar Aveli e Damão e Diu	marata	dadraense
Nagar Aveli	नगर हवेली Nagar Haveli	1779-1954	Dadrá e Nagar Aveli e Damão e Diu	marata	nagar-aveliense
Baçaim	वसई Vasai	1534-1739	Maarastra	marata	baçainense
Salsete	साल्सेट Salsette	1534-1737	Maarastra	marata	salsetense
Mombaim Bombaim	मुम्बई (Bombay) Mumbai	1534-1661	Maarastra	marata	mombainense bombainense
Chaul	कोर्लई Korlai	1521-1740	Maarastra	marata	chaulense
Goa	गोवा Goa	1509-1961	Goa	concanim	goês
Mangalor	मैंगलुरु (Mangalore) Mangaluru	1568-1763	Carnataca	canarim	mangalorense

Cananor	कन्नूर (Cannanore) Kannur	1502-1636	Querala	malaialim	cananorense
Calecute	कोड़िकोड (Calicut) Kozhikode	1510-1664	Querala	malaialim	calecutense
Cranganor	कोडुंगलूर (Cranganore) Kodungallur	1523-1661	Querala	malaialim	cranganorense
Cochim	कोच्चि (Cochin) Kochi	1502-1663	Querala	malaialim	cochinês
Coulão	कोल्लम (Quilon) Kollam	1502-1661	Querala	malaialim	coulamês
Tuticorim	तूतुकुड़ी (Tuticorin) Thoothukudi	1548-1658	Tâmil Nadu	tâmil	tuticorinês
Negapatão	नागपट्टिनम (Nagapatnam) Nagapattinam	1507-1657	Tâmil Nadu	tâmil	negapatamês
Meliapor	मयलापुर (Mylapore) Mayilaappur	1523-1749	Tâmil Nadu	tâmil	meliaporense

A observação desta lista de topónimos portugueses do Estado da Índia permite identificar alguns padrões bem definidos, sobretudo ao nível das **terminações**, mesmo para topónimos oriundos de diferentes substratos linguísticos indianos. Exemplos:

1. *terminações em «im» (ou «i»)*

A solução em «im» encontrada para representar o /iː/ final, **tónico**, longo e ligeiramente nasalado de muitos endónimos é uma das marcas mais típicas da toponímia indiana em português. Este som é representado em devanágari por ी (em sílaba) ou ई (isolado) e em AITS por ī.

É o caso, em Goa, de Pangim, mas também de Bicholim, Agaçaim, Sanquelim, Cortalim, Carambolim, Benastarim, etc. Em alguns casos, sobretudo nas Novas Conquistas (séc. XVIII), vingou a terminação em «i», como é o caso de Chaudi (e não Chaudim) ou de Satari (e não Satarim), noutros ainda coexistem variantes, como é o caso do nome da língua de Goa — concanim/concani. Também fora da zona do concanim há palavras terminadas em «im», representando um i longo, ligeiramente nasalado: Baçaim, Mombaim/Bombaim ou Tuticorim. Mas Aveli (e não Avelim), em Dadrá e Nagar Aveli. Outra terminação nasalada típica de Goa é o «ém» tónico, correspondente a ेम em devanágari, como em Perném, Quepém ou Sanguém.

2. *terminações em «ão»*

Tal como nas palavras derivadas do latim, a terminação «ão» tónica nasalada é a solução adotada para vários sons das línguas de origem, como seja /əm/ ou /ən/, representados por म ou न finais. É o caso de Coulão ou de Damão.

A terminação «gão», típica de Goa, é utilizada também para representar /gaːʊ/, como em Mormugão, Margão, Assagão, Corgão, Lamagão, Sirgão, Taleigão, etc., correspondendo ao elemento concanim *ganv* (comunidade/aldeia).

3. *terminações em «á»*

Outra terminação típica é em «á», para representar o /aː/ final tónico, correpondente ao devanágari ा (em sílaba) ou आ (isolado) e ao ā em AITS. Em Goa é o caso de Mapuçá, Pondá, Darbandorá, Candolá, Mopá, Majordá, Sirodá, Virnorá, etc. Também fora da zona do concanim: Dadrá.

4. *terminações em «or»*

A terminação tónica «or» (ोर em devanágari), como em Cananor, Cranganor, Mangalor, Meliapor, distingue-se da terminação «ore» tradicionalmente adotada em inglês.

Topónimos de estados e territórios da Índia

Com o declínio da presença portuguesa e com o papel desempenhado pelo inglês durante o período britânico e após a independência da Índia, o português deixou de ir buscar topónimos diretamente aos endónimos indianos, passando a adaptar ou simplesmente copiar fontes escritas, fundamentalmente em inglês. Um bom exemplo desta influência do inglês no português é a passagem de Himalaia a Himalaias. Os dicionários de língua portuguesa ainda registam Himalaia, mas os textos resultantes de traduções mais recentes transformaram a cordilheira num plural.

É também o caso da toponímia associada aos atuais estados e territórios da Índia, que nos chegou via o inglês britânico (por vezes adaptado do português) e, atualmente, via o inglês indiano (com a romanização AITS de endónimos das diferentes línguas indianas). Assim fazendo, e desconhecendo-se as regras ortográficas do inglês indiano, o resultado é uma pronúncia em português frequentemente mais afastada do endónimo do que seria o caso nos topónimos portugueses do primeiro contacto.

No entanto, atendendo à tradição de toponímia lusoindiana, pode chegar-se a exónimos portugueses, apresentados mais adiante em tabela, que se aproximem tanto quanto possível dos endónimos locais, com as habituais adaptações ortográficas e simplificações ao nível das terminações, a saber:

- distinção gráfica entre os sons /ʃ/e /t͡ʃ/, mesmo que já não produzida pela maioria dos locutores lusófonos:
  - /ʃ/ (श «x»): Caxemira, Pradexe, Xilongue, Ximelá, Orixá, etc.
  - /t͡ʃ/ (च «ch»): Aranuchal, Chandigar, Ranchi, Chenai, etc.
- «im»/«i» tónico final:
  - Pangim e Bombaim
  - Amaravati, Ranchi, Cavarati, Deli, Pondicherri (com «im» tónico final daria: Ranchim, Cavaratim, Delim, Pondicherrim)
  - «im» tónico final em Siquim será uma simplificação — Siquime estaria mais próximo da terminação िम
- «ão» tónico final:
  - सम «são» (Assão); रम «rão» (Mizorão); स्थान «stão» (Rajastão); मान «mão» (Andamão) — tal como Coulão e Negapatão ou Damão
  - mais recentemente, as terminações «ame» e «ã» têm conquistado lugar ao «ão» final na toponímia lusófona na Ásia — assim, encontram-se registos de Assame/Assã e Mizorame/Mizorã (cf. Vietname /Vietnã em vez de Vietnão(8) ou Bandar Seri Begauã/Begauane em vez de Bandar Seri Begauão)
- «á» tónico final:
  - ता «tá» (Calcutá), ना «ná» (Patná), ला «lá» (Ximelá, Agaratalá), मा «má» (Coimá), शा «xá» (Orixá), रा «rá» (Tripurá), णा «ná» (Harianá) — tal como Dadrá ou Mapuçá
- «or» tónico final:
  - लौर «lor» (Bangalor) — tal como Mangalor, Cananor, Cranganor ou Meliapor

O caso de Bombaim é emblemático e não seria totalmente descabido voltar a designar-se **Mombaim** em português. मुंबई (AITS: Mumbaī) é a capital do estado de Maarastra. O território, foi inicialmente designado Mombaim pelos portugueses por transcrição do marata(9). Foi posteriormente o primeiro posto britânico na Índia, fazendo parte do dote de Dona Catarina de Bragança no casamento com Carlos II de Inglaterra. Atualmente, em inglês indiano sem diacríticos escreve-se Mumbai, depois de em inglês britânico se ter chamado Bombay, a que corresponde o português Bombaim.

estado **pt** **en-IN**(10) **hi**(11)	**ISO 3166-2**	**gentílico**	**capital** **pt** **en-IN** **hi**	**IATE estado**
Andra Pradexe(12) Andhra Pradesh आन्ध्र प्रदेश	IN-AP	andra-pradexiano	**Amaravati** Amaravati अमरावती	3542046
Arunachal Pradexe(13) Arunachal Pradesh अरुणाचल प्रदेश	IN-AR	leste-pradexiano	**Itanagar**(14) Itanagar ईटानगर	3542047
Assão(15) Assam असम	IN-AS	assamês	**Dispur** Dispur दिसपुर	3590870
Bengala Ocidental West Bengal(16) पश्चिम बंगाल	IN-WB	oeste-bengalês	**Calcutá** Kolkata कोलकाता	
Biar Bihar बिहार	IN-BR	biarês	**Patná** Patna पटना	3531784
Carnataca(17) Karnataka कर्नाटक	IN-KA	carnataquês	**Bangalor** Bengaluru बंगलौर	3590212
Chatisgar(18) Chhattisgarh छत्तीसगढ़	IN-CT	chatisgarense	**Raipur** Raipur रायपुर	3590875
Goa Goa गोंय	IN-GA	goês	**Pangim** Panaji(19) पणजी(20)	
Guzerate(21) Gujarat गुजरात	IN-GJ	guzerate(22)	**Gandinagar**(23) Gandhinagar गाँधीनगर	3590871
Hariana Haryana हरियाणा	IN-HR	harianaês	**Chandigar**(24) Chandigarh चण्डीगढ़	3590881
Himachal Pradexe(25) Himachal Pradesh हिमाचल प्रदेश	IN-HP	himachal-pradexiano(26)	**Ximelá** Shimla शिमला	3542048
Jarcanda(27) Jharkhand झारखंड	IN-JH	jarcandês	**Ranchi** Ranchi राँची	3590876
Madia Pradexe(28) Madhya Pradesh मध्य प्रदेश	IN-MP	centro-pradexiano	**Bopal** Bhopal भोपाल	2230162
Maarastra(29) Maharashtra महाराष्ट्र	IN-MH	maarastriano	**Bombaim** Mumbai मुम्बई	3590213
Manipur Manipur मणिपुर	IN-MN	manipurês	**Infal** Imphal इम्फाल	

Megalaia[30] Meghalaya मेघालय	IN-ML	megalaico	**Xilongue**[31] Shillong शिलांग	3590877
Mizorão[32] Mizoram मिज़ोरम	IN-MZ	mizoramês	**Aizol** Aizawl अइज़ोल	3590872
Nagalândia[33] Nagaland नागालैण्ड	IN-NL	nagalandês	**Coimá** Kohima कोहिमा	3590885
Orixá[34] Odisha[35] ओडिशा	IN-OR	orixaês	**Bubanesvar** Bhubaneswar भुवनेश्वर[36]	3590878
Penjabe[37] Punjab पंजाब	IN-PB	penjabês	**Chandigar** Chandigarh चण्डीगढ़	2224825
Querala Kerala केरल	IN-KL	queralita	**Tiruvanantapura** Thiruvananthapuram तिरुवनन्तपुरम	3590211
Rajastão[38] Rajasthan राजस्थान	IN-RJ	rajastanês	**Jaipur**[39] Jaipur जयपुर	3590886
Siquim Sikkim सिक्किम	IN-SK	siquimês	**Ganguetoque** Gangtok गान्तोक	1325905
Tâmil Nadu[40] Tamil Nadu तमिल नाडु	IN-TN	tâmil	**Chenai** Chennai चेन्नई	3590884
Telanganá Telangana तेलंगाना	IN-TG	telugo	**Haiderabade** Hyderabad हैदराबाद	3590882
Tripurá Tripura त्रिपुरा	IN-TR	tripuraês	**Agartalá** Agartala अगरतला	3590883
Utar Pradexe[41] Uttar Pradesh उत्तर प्रदेश	IN-UP	norte-pradexiano	**Lucnau**[42] Lucknow लखनऊ	346399
Utaracanda[43] Uttarakhand उत्तराखण्ड	IN-UT	norte-candês	**Deradum** Dehradun देहरादून	3590873
territórios da União **pt** **en-IN** **hi**	**ISO 3166-2**	**gentílico**	**capital** **pt** **en-IN** **hi**	**IATE território**
Andamão e Nicobar Andaman and Nicobar Islands अण्डमान और निकोबार द्वीपसमूह	IN-NA	andamanês nicobarês	**Porto Blair** Port Blair पोर्ट ब्लेयर	
Chandigar[44] Chandigarh चण्डीगढ़	IN-CH	chandigariense	**Chandigar** Chandigarh चण्डीगढ़	3590880
Dadrá e Nagar Aveli e Damão e Diu[45] Dadra and Nagar Haveli and Daman and Diu दादरा और नगर हवेली और दमन और दीव	IN-DD	dadraense nagar-aveliense damanense diuense[46]	**Damão** Daman दमन	

Jamu e Caxemira[47] Jammu and Kashmir जम्मू और कश्मीर	IN-JK	caxemiriano	**Serinagar**[48] Srinagar श्रीनगर	387105
Ladaque Ladakh लद्दाख़	IN-LA	ladaquiano	**Lé** Leh लेह	3590879
Laquedivas[49] Lakshadweep लक्षद्वीप	IN-LD	laquedivo	**Cavarati** Kavaratti कवरत्ती	3590874
Deli[50] Delhi दिल्ली	IN-DL	deliense	**Deli** Delhi दिल्ली	
Pondicherri[51] Puducherry पुदुच्चेरी	IN-PY	pondicherriano	**Pondicherri** Puducherry पुदुच्चेरी	

Alguns elementos etimológicos são reconhecíveis em vários topónimos, quer dos nomes dos estados quer dos nomes das cidades. Exemplos:

- **pradexe** (प्रदेश) — província/estado (incluindo **dexe**[52] (देश) — país —, como em Bangladexe):
  - Andra Pradexe — Província dos Andras (povo)
  - Arunachal Pradexe — Província do Levante
  - Himachal Pradexe — Província Nevada (cf. Himalaia — Morada da Neve)
  - Madia Pradexe — Província Central
  - Utar Pradexe — Província do Norte (cf. Utaracanda — Terra do Norte)
- **canda** (खण्ड) — terra:
  - Jarcanda — Terra das Florestas
  - Utaracanda — Terra do Norte
- **laia** (लय) — morada
  - Megalaia — Morada das Nuvens
- **gar** (गर) — forte
  - Chandigar — Forte de Chandi
  - Chatisgar — Trinta e Seis Fortes
- **nagar** (नगर) — cidade
  - Gandinagar — Cidade de Gândi
  - Itanagar — Cidade de Tijolo
  - Serinagar — Cidade Resplandecente (cf. Seri Lanca)
- **pur/por/pura** (पुर) — cidade, terra
  - Jaipur — Cidade da Vitória (cf. Jaiapura, Indonésia, e Seri Jaiavardenapura, Seri Lanca)
  - Raipur — Cidade de Rai (cf. Singapura— Cidade do Leão)
  - Manipur — Terra das Joias

Paulo.Correia@ec.europa.eu

Anexo: Devanágari — vogais, vogais diacríticas e consoantes

		AFI[53]/ AITS[54]		**AFI/ AITS**		**AFI/ AITS**		**AFI/ AITS**		**AFI/ AITS**		**AFI/ AITS**		**AFI/ AITS**
अ		ə / a	क	k / k	च	t͡ʃ / c	ट	ʈ / t	त	t̪ / t	प	p / p	य	j / y
आ	◌ा	aː / ā	ख	kʰ / kh	छ	t͡ʃʰ / ch	ठ	ʈʰ / th	थ	t̪ʰ / th	फ	pʰ / ph	र	r / r
इ	ि◌	ɪ / i	ख़	x	—	—	—	—	—	—	फ़	f	ल	l / l
ई	◌ी	iː / ī	ग	g / g	ज	d͡ʒ / j	ड	ɖ / d	द	d̪ / d	ब	b / b	व	ʋ / v
उ	◌ु	ʊ / u	ग़	ɣ	ज़	z	ड़	ɖɽ	—	—	—	—	—	—
ऊ	◌ू	uː / ū	घ	gʱ / gh	झ	ɲ / jh	ढ	ɖʱ / dh	ध	d̪ʱ / dh	भ	bʱ / bh	श	ʃ / ś
ए	◌े	eː / e	—	—	—	—	ढ़	ɽʱ	—	—	—	—	ष	ʂ / ṣ

ऐ	◌ै	ɛː / ai	ङ	ŋ / ṅ	ञ	d͡ʒʱ / ñ	ण	ɳ / ṇ	न	n / n	म	m / m	स	s / s
ओ	◌ो	oː / o	—	—	—	—	—	—	—	—	—	—	ह	h / h
औ	◌ौ	ɔː / au	—	—	—	—	—	—	—	—	—	—	—	—

---

(1) Bocarro, A., *Livro das Plantas de Todas as Fortalezas, Cidades e Povoações do Estado da Índia Oriental*, Biblioteca Nacional Digital, http://purl.pt/27184/5/P168.html, http://purl.pt/27184/5/P169.html.
(2) Não confundir com Salcete, taluca (concelho) de Goa.
(3) Por decalque do inglês Bay of Bengala, o golfo de Bengala é por vezes referido como baía de Bengala.
(4) O termo canarim também era utilizado, de forma depreciativa, para designar os naturais da Índia Portuguesa, sobretudo de Goa.
(5) Governo da Índia, *Constituição da Índia*, https://www.india.gov.in/sites/upload_files/npi/files/coi_part_full.pdf.
(6) As pronúncias podem ser ouvidas, por exemplo, em Google Tradutor, https://translate.google.com/?um=1&ie=UTF-8&hl=pt&client=tw-ob#hi/pt.
(7) Wikipédia, *Alfabeto Internacional para a Transliteração de Sânscrito*, https://pt.wikipedia.org/wiki/Alfabeto_internacional_para_a_translitera%C3%A7%C3%A3o_de_s%C3%A2nscrito.
(8) Cf. Teixeira, M., *Carta do Vietnão*, http://www.library.gov.mo/ManuelTC/ppart2ct.htm#165.
(9) Como se pode ouvir no *Google Tradutor*: https://translate.google.com/?rlz=1C1CHZL_pt-PTPT745PT745&um=1&ie=UTF-8&hl=pt&client=tw-ob#mr/pt/%E0%A4%AE%E0%A5%81%E0%A4%82%E0%A4%AC%E0%A4%88.
(10) Governo da Índia, *States of India*, https://www.india.gov.in/india-glance/states-india.
(11) Governo da Índia, भारत के राज्य, https://www.india.gov.in/hi/india-glance/states-india.
(12) Província dos Andras (povo).
(13) Província do Levante.
(14) Itanagar, cidade de Ita.
(15) *Dicionário Lello*: Assame ou Assão; *Infopédia*: Assame; *Dicionário Aurélio*: Assã.
(16) Fazendo fé no jornal *The Hindu*, em língua inglesa, foi proposto que West Bengal passe a chamar-se Paschim Banga, que significa mais precisamente Bengala Ocidental! Haverá necessidade de o português embarcar neste tipo de mudanças? «West Bengal may be renamed PaschimBanga», *The Hindi*, 19.8.2011, https://www.thehindu.com/news/national/west-bengal-may-be-renamed-paschimbanga/article2373155.ece.
(17) «"Posso assegurar-vos que Portugal é um país que receberá de braços abertos caso entendam ali investir, trabalhar ou viver"», afirmou o Primeiro-Ministro António Costa no seu discurso na sessão de abertura da convenção mundial da diáspora indiana, que reuniu milhares de representantes das comunidades indianas expatriadas, em Bangalore, capital do Estado de Carnataca.», «Portugal acolheu muitos indianos e "receberá de braços abertos" os que queiram investir, trabalhar ou viver», *República Portuguesa: XXI Governo Constitucional*, https://www.portugal.gov.pt/pt/gc21/comunicacao/noticia?i=20170108-pm-india.
(18) Trinta e seis fortes — *chatis* (36).
(19) Antes Panjim.
(20) Antes पंजिम.
(21) *Infopédia* e *Priberam*: Guzerate; *Priberam*: Guzarate.
(22) Ou guzarates, Wikisource, *Os Lvsíadas/X*, https://pt.wikisource.org/wiki/Os_Lvsiadas/X.
(23) Gandinagar, cidade de Gândi. Cidade construída nos anos 60 do século XX em homenagem ao guzerate Maatma Gândi.
(24) Chandigar é a capital de dois estados — Harianá e Penjabe —, situando-se num território federal.
(25) Província Nevada.
(26) Nivipradexiano.
(27) Terra das Florestas.
(28) Província Central.
(29) Grande Nação (cf. marajá — grande rei; maatma — grande mestre).
(30) Morada das Nuvens (cf. Himalaia — Morada das Neves — e himalaico).
(31) «Estes três irmãos conseguiram bem mais de 200 assinaturas nos primeiros dois meses em Xilongue.», «Índia», *Anuário das Testemunhas de Jeová*, 1978, https://wol.jw.org/pt/wol/d/r5/lp-t/301977003.
(32) Terra dos Montanheses. Separou-se de Assão em 1972. Variantes: Mizorame ou Mizorã.
«A sudeste de Manipur acha-se Mizorã e Tripura.», «Índia», *Anuário das Testemunhas de Jeová*, 1978, https://wol.jw.org/pt/wol/d/r5/lp-t/301977003.
(33) Terra dos Nagas.
(34) Orixá, Porto Editora, *Dicionário Infopédia da Língua Portuguesa: oriá*, https://www.infopedia.pt/dicionarios/lingua-portuguesa/oriá.
«Ha tambem nesta cidade muita soma de almizcre muito milhor que o da China, que tambem se leva para Martavão & Peguu, onde os nosso o compraõ para de veniaga o levarem a Narsinga, Orixaa, & Masulepatão.», Pinto, F. M., *Peregrinação: Capítulo CXXX*, GoogleBooks, https://books.google.be/books?id=QBBFAQAAMAAJ.
(35) Antes Orissa.
(36) श्र — ligatura de श e र.
(37) Terra dos Cinco Rios. Ver «penjábico», Lexikon, *Dicionário Aulete Digital*, http://www.aulete.com.br/penjábico.
(38) País dos Rajás, País dos Reis.
(39) Cidade da Vitória. Cf. Jaiapura (Iriã Jaia, Indonésia)

(40) Terra dos Tâmiles.
(41) Província do Norte.
(42) «Russell e seu grupo também prepararam a base para a expansão futura, por proferirem discursos por toda a Índia, na cidade religiosa de Benares, na histórica cidade de Lucnau e em Trivandrum, Cotaracara, Nagercoil, Purã e Vizagapatã, bem como nos portos comerciais de Calcutá e Bombaim.», «Índia», *Anuário das Testemunhas de Jeová*, 1978, https://wol.jw.org/pt/wol/d/r5/lp-t/301977003#h=2.
(43) Terra do Norte. Até 2006: Utaranchal.
(44) Forte de Chandi.
(45) Resultou da fusão em 26 de janeiro de 2020 dos territórios de Dadrá e Nagar Aveli (com capital em Silvassa, antiga Vila de Paço d'Arcos) e de Damão e Diu.
(46) «(…) a identidade própria dos goeses provém principalmente da sua ancestralidade concani, a dos damanenses e diuenses do seu substrato guzerate e a dos dadraenses e nagar-avelienses da sua base marata.», Fernandes, M. L. F. C., *Os «Encontros dos Goeses da Beira e Seus Amigos» na Grande Lisboa: Património Identitário e Sua Reconstrução Cultural Pós-Colonial*, http://hdl.handle.net/10362/4241.
(47) Estado até 2019.
(48) Cidade notável, cidade resplandecente (cf. Seri Lanca).
(49) Cem mil ilhas. Ver «laque», Lexikon, *Dicionário Aulete Digital*, http://www.aulete.com.br/laque.
(50) Território da Capital Nacional. «Déli», *Vocabulário Ortográfico Comum da Língua Portuguesa*, https://voc.cplp.org/index.php?action=toponyms&act=details&id=TER.142.034.IN.
(51) Quatro enclaves costeiros, que constituíam a antiga Índia Francesa: Pondichéry (ou Puducherry, antes Pondicherry), Karikal (Karaikal), Mahé (Mahe) e Yanaon (Yanam).
(52) Em Goa a solução adotada para a taluca (concelho) de Bardez (Doze Terras) foi a terminação em «dez» em vez de «dexe».
(53) Wikipedia, *Hindustani Orthography*, https://en.wikipedia.org/wiki/Hindustani_orthography.
(54) Wikipedia, *International Alphabet of Sanskrit Transliteration*, https://en.wikipedia.org/wiki/International_Alphabet_of_Sanskrit_Transliteration.

Um aparte à parte (IV)

Jorge Madeira Mendes
Antigo funcionário da Direção-Geral da Tradução — Comissão Europeia

Não diga «mudança *climatérica*», «condições *climatéricas*». Diga «alterações ***climáticas***», «condições ***climáticas***» (ou, melhor até, «condições ***meteorológicas***»).

Explicação:
O adjetivo «climático(s)/a(s)» significa «relativo ao clima; climatológico». Por sua vez, o adjetivo «climatérico(s)/a(s)», embora possa também significar «climatológico», tem a aceção principal de «relativo ao climatério», sendo «climatério» o período correspondente à menopausa na mulher ou ao declínio sexual no homem[1]. Para evitar ambiguidades, parece, pois, preferível reservarmos «climático(s)/a(s)» para o domínio do clima: além de ortograficamente mais simples, é, neste domínio, um termo mais consagrado do que «climatérico(s)/a(s)».

jorge.mendes909@gmail.com

(1) Cf. Volume 1 (Léxico Comum) do *Dicionário Enciclopédico Koogan-Larousse-Seleções*, Seleções do Reader's Digest, Lisboa, Rio de Janeiro, Nova Iorque, 1980, 3.ª ed.

Um aparte à parte (V)

Jorge Madeira Mendes
Antigo funcionário da Direção-Geral da Tradução — Comissão Europeia

Não diga /*ÓNU*/. Diga /*ONÚ*/.

Explicação:

A sigla O.N.U., designativa da Organização das Nações Unidas, deve ler-se como as palavras portuguesas *tabu*, *caju* ou *peru*, que, apesar de não regerem acento gráfico no *u* final, são palavras agudas, isto é, a sua sílaba tónica é a última (as grafias *tabú*, *cajú* e *perú*, conspícuas nos órgãos de informação escrita, estão erradas, por redundância grosseira).

Isto porque as convenções da ortografia portuguesa estabelecem que a vogal *u*, quando no final da palavra e não antecedida de outra vogal (portanto, quando não suscetível de se integrar num ditongo), «puxa» automaticamente o acento para si, pelo que prescinde da expressão gráfica desse acento.

Friso a condição de a vogal terminal *u* não ser antecedida de outra vogal, porque nesse caso formaria um ditongo com ela (p. ex., *pau*, *balandrau*, *cacau*). Em situações nas quais fosse necessário desfazer tal ditongo, o acento gráfico teria efetivamente de figurar (p. ex., *baú*). Do que estamos a falar é de *u* final não precedido de outra vogal: em *tabu*, o *u* final é precedido da consoante *b*, em *caju* é-o da consoante *j*, em *peru* é-o da consoante *r*... em ONU é-o da consoante *n*. Portanto, todas estas palavras se pronunciam acentuando o *u* final, sem exprimir graficamente essa acentuação.

Uma consequência perversa desta confusão no tocante à tonicidade do *u* final é a crescente tendência do meio culinário para pronunciar a palavra *menu* como /*meno*/. A «lógica» deste atropelo parece ser a seguinte: se *menu* se escreve sem acento no *u*, então a sílaba tónica deve ser a primeira.

Voltemos ao princípio: não diga /*ÓNU*/, diga /*ONÚ*/!

jorge.mendes909@gmail.com

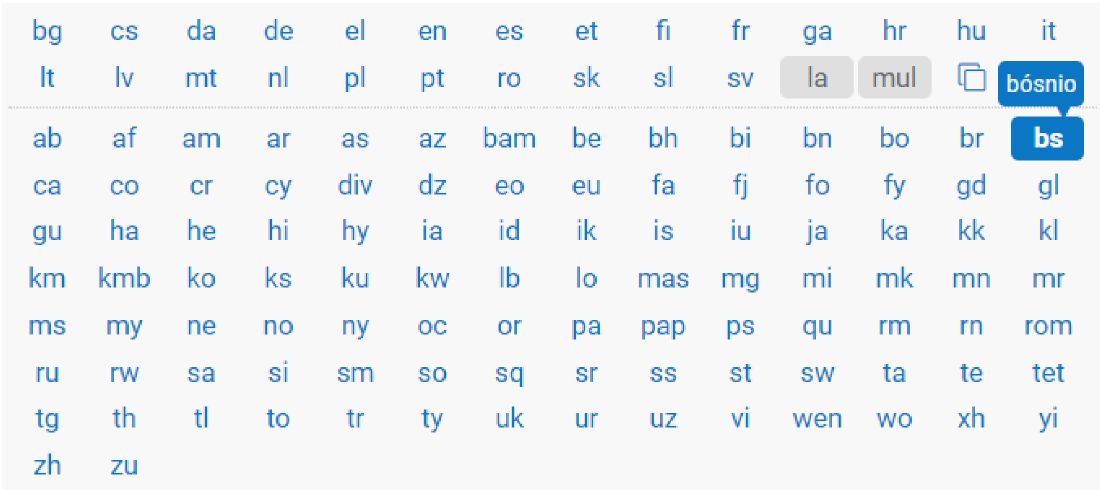
As línguas da IATE — notas de tradutor

Paulo Correia
Direção-Geral da Tradução — Comissão Europeia

De longa data que, no quadro do *Código de Redação Interinstitucional* (CRI) e da base terminológica IATE, o Grupo Interinstitucional de Terminologia Portuguesa (GITP) se tem debruçado sobre topónimos, gentílicos e nomes de moedas de países e territórios de todo o mundo (anexo A5 do CRI(1)). Porém, o GITP nunca tinha tido oportunidade de desenvolver um trabalho mais sistemático sobre os **glossónimos** que fosse além das designações das línguas oficiais e de trabalho da União Europeia (anexo A8 do CRI(2)).

A localização da interface portuguesa da base IATE foi essa oportunidade. Tratava-se de dar resposta à pergunta colocada pelo utilizador da interface portuguesa da IATE ao passar o cursor do rato sobre a

lista das línguas de partida e de chegada: qual é o nome das línguas representadas pelos respetivos códigos **ISO 639-1** e **ISO 639-2**[3]?



A base terminológica IATE contém terminologia nas 24 línguas oficiais da União Europeia. Essas línguas são maioritariamente da grande família indo-europeia, mas há também o estónio, o finlandês e o húngaro — línguas uralo-altaicas — e o maltês — língua semítica. Também o latim (**la**) está presente desde o início das bases terminológicas que antecederam a base IATE, juntando-se-lhe a «língua» multilingue (**mul**), usada para termos partilhados pelo conjunto das línguas (exemplo: os próprios códigos ISO das línguas).

Além destas, a IATE inclui as línguas oficiais dos países candidatos ou potenciais candidatos a uma futura adesão à União Europeia[4] — com exceção do montenegrino[5] —, assim como duas línguas artificiais — o esperanto (**eo**) e a interlíngua (**ia**)[6] — e outras largas dezenas de línguas de todo o mundo, das maiores às menores, de várias macrofamílias, famílias e ramos linguísticos.

Levantamento de glossónimos

Para acompanhar o trabalho de localização da interface IATE foi feito um levantamento dos glossónimos registados em **fontes lexicográficas** do português[7] — dicionários e, quando necessário, em vocabulários ortográficos —, como:

PE	Infopédia[8]
DP	Dicionário Priberam[9]
DE	Dicionário Estraviz[10]
AD	Aulete Digital[11]
DM	Dicionário Michaelis[12]
DH	Dicionário Houaiss[13]
DA	Dicionário Aurélio[14]
VOP	Vocabulário Ortográfico do Português[15]
VOAC	Vocabulário Ortográfico da Língua Portuguesa (Academia das Ciências de Lisboa)[16]
VOLP	Vocabulário Ortográfico da Língua Portuguesa (Academia Brasileira de Letras)[17]

Com a exceção das línguas esquimó-aleútes, praticamente todas as línguas presentes na IATE têm já nomes portugueses registados nos dicionários e vocabulários da língua portuguesa.

Porém, a análise do uso dos glossónimos menos familiares em **traduções portuguesas** (incluindo nas memórias de tradução interinstitucionais Euramis) revela uma certa tendência para a não tradução ou para o decalque dos termos e ortografia utilizados pelo francês e ou, mais modernamente, pelo inglês, ignorando, por vezes, os nomes nas fontes lexicográficas lusófonas e nas próprias línguas. Exemplos:

pachto (fr: *pachto*) / Pashto (en: *Pashto*) / pastó	ps: *paẍtō*
tâmul (fr: *tamoul*) / tâmil (en: *Tamil*)	ta: *tamiḻ*
feroês (fr: *féroïen*) / faroês (en: *Faroese*)	fo: *føroyskt*

Por vezes, ao decalcar-se e não se adaptar completamente a ortografia, o português afasta-se da pronúncia na língua local (e mesmo da pronúncia em inglês). Exemplos:

s/ss	issixossa ou, simplesmente, xossa refletem melhor do que isixosa a pronúncia de *isiXhosa* (xh)
r/rr	quiniarruanda ou, simplesmente, ruanda refletem melhor do que quiniaruanda a pronúncia de *Ikinyarwanda* (rw)

Nas **tabelas** em anexo a este artigo são registados os glossónimos na própria língua, as diferentes designações em português recolhidas nas fontes lexicográficas e as designações em inglês e francês adotadas na norma ISO 639. Apresentam-se igualmente **notas de tradutor**, procurando esclarecer, quando necessário, a opção — indicada a negrito — por um glossónimo em português em detrimento de outros igualmente registados.

Por facilidade de análise, essas tabelas agrupam as línguas por macrofamílias, famílias e ramos. Esta classificação das línguas nem sempre é consensual entre os linguistas:

1. LÍNGUAS EURASIÁTICAS
    - 1.1 línguas indo-europeias
        - — indo-iranianas, ítalo-célticas, germânicas, balto-eslavas, albanesa, arménia, helénica
    - 1.2 línguas uralo-altaicas
        - — urálicas, altaicas, coreana, japonesa
    - 1.3 línguas esquimó-aleútes
2. LÍNGUAS AFRO-ASIÁTICAS
    - — semíticas, cuxíticas, chádicas
3. LÍNGUAS SUL-SARIANAS
    - 3.1 línguas nilo-sarianas
    - 3.2 línguas nigero-congolesas
4. LÍNGUAS ÁUSTRICAS
    - 4.1 línguas tai
    - 4.2 línguas austro-asiáticas
    - 4.3 línguas austranésias
5. LÍNGUAS AMERÍNDIAS
6. LÍNGUAS DENE-CAUCASIANAS
    - 6.1 língua basca
    - 6.2 línguas caucasianas
    - 6.3 línguas dravídicas
    - 6.4 línguas sino-tibetanas

É também sempre bom recordar que, em português, contrariamente ao inglês, os nomes das línguas são escritos com **minúscula inicial**, mesmo no caso de línguas cujo aportuguesamento ainda não está registado, só se utilizando maiúscula inicial quando de trata do nome de disciplinas escolares. Exemplo:

Inglês II, Francês III

Paulo.Correia@ec.europa.eu

Anexo: Macrofamílias, famílias e ramos linguísticos presentes na base IATE

1. LÍNGUAS EURASIÁTICAS

As línguas eurasiáticas constituem uma macrofamília, proposta por alguns autores, com línguas tão diferentes como as línguas indo-europeias, as línguas uralo-altaicas ou mesmo as línguas esquimó-aleútes.

1.1 *Línguas indo-europeias*

A grande família indo-europeia subdivide-se em vários ramos.

1.1.1 *Línguas indo-iranianas*

Ramo mais oriental da família indo-europeia, subdividido nos sub-ramos indo-ariano e iraniano, com línguas faladas da Anatólia oriental ao sul do Seri Lanca. Neste grupo encontram-se dois importantíssimos diassistemas(18): urdu-hindi (ou hindustani) e persa-tajique.

ISO 639	endónimo (19)	pt	en	fr	IATE
as	অসমীয়া (asamīyā)	**assamês** PE, AD, DM, DH, DA assami AD, DM, DH	Assamese	assamais	1451547
bh	बिहारी (bihārī)	**biari** AD, DM, DH, DA; (ver nota 2)	Bihari	bihari	282196
bn	বাংলা (bānlā)	**bengali** PE, DP, DE, AD, DM, DH, DA bengalim VOLP	Bengali	bengali	282198
div	ދިވެހި (divehi) ދިވެހިބަސް (divehibas)	**maldivano** PE, DP, DE, AD; (ver nota 3) maldiviano PE, DE maldívio PE, DM, DH maldivo DP maldivense PE, DM *divehi* DH diveí	Divehi Dhivehi Maldivian	maldivien	282285
fa	فارسی (fārsī)	**persa** PE, DP, DE, AD, DM, DH, DA farsi PE, DM, DA pársi PE, DP, DH, DA parse PE, DP, DH, DA	Persian	persan	282206
gu	ગુજરાતી (gujarātī)	**guzerate** PE, DP, DE, AD, DM, DH, DA guzeráti DA guzarate PE, DP, DE, AD, DM, DH, DA gujaráti AD gujarate VOLP	Gujarati	goudjrati	282211
hi	हिन्दी (hindī)	**hindi** PE, DP, AD, DA híndi PE, DP, DE, DM indi AD índi DM	Hindi	hindi	282213
ks	कश्मीरी کأشُر (kaśmīrī)	**caxemiriano** PE caxemira PE, DP, DH caxemíri DA	Kashmiri	kashmiri	282227
ku	kurdî کوردی	**curdo** PE, DP, DE, AD, DM, DH, DA	Kurdish	kurde	282228
mr	मराठी (marāṭhī)	**marata** PE, DP, DE, AD, DM, DH, DA marati DM, DH, DA	Marathi	marathe	282237
ne	नेपाली (nēpālī)	**nepalês** PE, DP, AD, DM, DH, DA nepali PE, DE, AD, DM, DH, DA	Nepali	népalais	282241
or	ଓଡ଼ିଆ (oṛiā)	**oriá** PE, DP, DE, AD, DM, DH, DA	Oriya	oriya	282244

pa	ਪੰਜਾਬੀ پنجابی (pajābī)	**panjabi** [PE, AD, DM, DH, DA; (ver nota 1)] **panjábi** [PE] penjabi [VOLP] penjabês [VOP] punjabi [PE, DM, DH, DA]	Panjabi Punjabi	pendjabi	282245
ps	پښتو (paxtō)	**pastó** [PE, DP, DE, AD,] pachto [DH] afegão [PE, DM] afegane [PE, AD, DM]	Pushto Pashto	pachto	282247
rom	romani čhib	**romani** [PE, DP. DE, AD, DM] românî [AD] romanho [PE, DP] cigano [PE, DP, AD, DM]	Romany	tsigane	282283
sa	संस्कृतम् (saṃskṛtam)	**sânscrito** [PE, DP, DE, AD, DM, DH, DA]	Sanskrit	sanskrit	282253
si	සිංහල (siṁhala)	**singalês** [AD; (ver nota 4)] **cingalês** [PE, DP, DE, AD, DM, DH, DA] cingalá [DP, AD]	Sinhala Sinhalese	singhalais	282255
tg	тоҷикӣ toçikī تاجیکی	**tajique** [PE, DP, DE, DH] tadjique [DP, AD, DM, DH, DA]	Tajik	tadjik	282267
ur	اُردُو (urdū)	**urdu** [PE, DP, DE, AD, DM, DH, DA]	Urdu	ourdou	282274

<u>Nota 1</u>: O persa manteve-se durante vários séculos como língua de prestígio no subcontinente indiano, sendo a principal língua da administração, primeiro no sultanato de Deli e posteriormente no império mogol, até ser suplantado pelo inglês no século XIX. Não é, assim, de estranhar que em vários glossónimos se verifique a influência do persa, refletida na terminação em «i», indicativa de «língua de». Exemplos:

hindi — a língua da Índia
bengali — a língua de Bengala
panjabi — a língua do Penjabe
nepali — a língua do Nepal
biari — a língua de Biar
romani — a língua dos romes

Tanto quanto é possível apurar consultando transcrições fonéticas de dicionários ingleses e escutando a pronúncia local, as palavras nas línguas locais tendem a ser agudas enquanto as palavras inglesas tendem a ser graves. Porém, embora não incluído na presente lista, há o caso da língua de Goa, que nos está muito próximo: o **concanim** ou **concani** (कोंकणी) — etimologicamente, a língua do Concão. O facto de o português ter adotado as **formas agudas** concanim/concani poderá servir de modelo para os restantes casos.

<u>Nota 2</u>: O **biari** inclui várias línguas, entre as quais: boiapuri (**bho**) — भोजपुरी/*bhōjapurī*; magaí (**mah**) — मगही/*magahī*; maitili (**mai**) — मैथिली/*maithilī*.

<u>Nota 3</u>: O **maldivano** é a língua das Maldivas. **Diveí** é o aportuguesamento do endónimo *devahi* (etimologicamente, insular).

<u>Nota 4</u>: A forma **singalês** parece preferível a cingalês, tendo em conta a etimologia, *siṃha*, leão em sânscrito (cf. Singapura, que substituiu Cingapura).

1.1.2 *<u>Línguas ítalo-célticas</u>*

O ramo ítalo-céltico da família de línguas indo-europeias subdivide-se nos sub-ramos itálico e céltico. Algumas destas línguas formam diassistemas: galego-português e occitano-catalão. Inclui também crioulos: hatiano — crioulo francês — e papiamento (de papear/papiar) — crioulo luso-espanhol.

ISO 639	endónimo	pt	en	fr	IATE
br	brezhoneg	**bretão** [PE, DP, DE, AD, DM]	Breton	breton	282200
ca	català	**catalão** [PE, DP, DE, AD, DM]	Catalan	catalan	282201
co	corsu lingua corsa	**corso** [DP] córsico [PE]	Corsican	corse	146970
cy	Cymraeg	**galês** [PE, DP, DE, AD]	Welsh	gallois	146971
es	español castillano	**espanhol** [CRI] castelhano [PE, DE, AD, DH, DA] castelão [DP, DE, DH]	Spanish Castilian	espagnol castillan	3529022
fr	français	**francês** [CRI]	French	français	3529020
ga	Gaeilge	**irlandês** [CRI] gaélico irlandês	Irish	irlandais	282208
gd	Gàidhlig	**escocês** [PE, DP, DE, AD, DM; (ver nota 1)] **gaélico escocês**	Gaelic Scottish Gaelic	gaélique gaélique écossais	282209
gl	galego	**galego** [PE, DP, DE, AD, DM]	Galician	galicien	282210
ht [cr[20]]	kreyòl ayisyen	**haitiano** [(ver nota 3)] **crioulo** crioulo haitiano	Haitian Haitian Creole	haïtien créole haïtien	1451571
it	italiano	**italiano** [CRI]	Italian	italien	3529028
kw	Kernewek	**córnico** [PE, DP, AD] cornualhês	Cornish	cornique	287221
oc	occitan	**occitano** [PE, DP, DE, DM; (ver nota 4)] occitânico [PE, DP, DE, AD, DM]	Occitan	occitan	282243
pap	Papiamentu	**papiamento** [PE, DP, DE, AD, DM] papeamento [AD, DM]	Papiamento	papiamento	1451607
pt	português	**português** [CRI]	Portuguese	portugais	3529026
rm	rumantsch	**romanche** [PE, DE, DM; (ver nota 5)] grisão [PE, DE]	Romansh	romanche	282249
ro	română	**romeno** [CRI] valáquio [PE, DP, DE, AD, DM]	Romanian	roumain	282250

Nota 1: As línguas célticas insulares subdividem-se em gaélico e britónico:

- **gaélico**, que engloba: irlandês (**ga**); escocês[21] (**gd**); manês (**gv**);
- **britónico**, que engloba: galês (**cy**); bretão (**br**); córnico (**kw**).

Nota 2: Algumas línguas adquirem nomes diferentes em diferentes entidades territoriais:

- **valenciano** (*valencià*) — designação da língua catalã na Comunidade Valenciana;
- **moldavo** (*limba moldovenească*) — designação da língua romena na República da Moldávia[22].

Nota 3: O **haitiano** é um crioulo francês. No entanto, a interface IATE não utiliza aqui o código ISO 639 (**ht**), mas uma abreviatura de crioulo (**cr**). O código ISO 639 **cr** corresponde à língua **cri** (ᓀᐦᐃᔭᐍᐏᐣ), língua ameríndia.

Nota 4: O termo **provençal** que já foi utilizado para designar as diferentes variedades do **occitano**, é agora utilizado apenas para designar a variedade do occitano falada na Provença. Variedades do occitano: auvérnio, limusino, vivaro-alpino, provençal, niçardo, languedociano, gascão.

Nota 5: Os termos **reto-romano**, **reto-romance**, **reto-românico** ou **rético** englobam o conjunto de línguas e dialetos românicos falados na região alpina central e oriental (sobretudo Suíça e Itália), correspondente à antiga Récia romana. Podem considerar-se três subgrupos: romanche (**rm**); ladino[23] (**lld**); friulano (**fur**).

O **romanche** (*rumantsch*) engloba o conjunto de línguas e dialetos reto-românicos falados no cantão dos Grisões, na Suíça. O romanche grisão (*rumantsch grischun*) é a norma unificada e artificial das

diferentes variantes do romanche (sursilvano, sutsilvano, surmirano, púter e valáder), criada como uma solução de compromisso, utilizada quase exclusivamente para fins administrativos.

1.1.3 Línguas germânicas

Ramo da família indo-europeia, que se subdivide em línguas germânicas nórdicas e germânicas ocidentais. Incluem-se o crioulo bislamá (Vanuatu) e o iídixe (dos judeus asquenazes).

ISO 639	endónimo	pt	en	fr	IATE
af	Afrikaans	**africanse** PE, DP; (ver nota 1) **africânder** PE, DP, DE, AD, DM africâner PE, DP, DE, AD, DM africanês PE, AD *afrikaans* DP, DM	Afrikaans	afrikaans	918200
bi	Bislama	**bislamá** PE	Bislama	bichlamar	282197
da	dansk	**dinamarquês** CRI	Danish	danois	
de	Deutsch	**alemão** CRI	German	allemand	
en	English	**inglês** CRI	English	anglais	
fo	føroyskt	**faroês** DP faroense DP feroês PE, DP, AD, DM feroico PE, DP, DE, AD, DM	Faroese	féroïen	282207
fy	Frysk	**frísio** PE, DP, DE, DM; (ver nota 2) frisão PE, DP, DM frisiano DP, DM frísico PE, DE, AD	Western Frisian	frison occidental	146973
is	íslenska	**islandês** PE, DP, DE, AD, DM	Icelandic	islandais	282218
lb	Lëtzebuergesch	**luxemburguês** DP	Luxembourgish Letzeburgesch	luxembourgeois	282282
nl	Nederlands	**neerlandês** CRI; (ver nota 3)	Dutch	néerlandais	3529041
no	norsk	**norueguês** PE, DP, DE, AD, DM; (ver nota 4) norueguense PE, DP, DE	Norwegian	norvégien	282242
sv	svenska	**sueco** CRI	Swedish	suédois	
yi	ייִדיש (yídish)	**iídiche** (ver nota 5) **iídiche** PE, DP, DE, AD, DM ídixe ídiche DP, DE, AD, DM *yiddish* DP judeo-alemão AD, DM judeu-alemão AD, DM	Yiddish	yiddish	282221

Nota 1: O aportuguesamento **africanse** parece preferível a africânder para representar o nome da língua falada, isso sim, pelos africânderes e mestiços do Cabo.

Nota 2: O frísio (**fy**) inclui três variantes: frísio do norte (**frr**); frísio oriental (**frs**); frísio ocidental (**fry**).

Nota 3: O neerlandês pode adquirir nomes diferentes em diferentes entidades territoriais:

- **flamengo** (*Vlaams*) — designação da língua neerlandesa na Bélgica;
- **holandês** — designação de variante do neerlandês; designação inexata da língua neerlandesa.

O limburguês (**li**) é uma língua aparentada, mas distinta do neerlandês.

Nota 4: O **norueguês** (**no**) inclui duas variantes: dano-norueguês (**nb**) — *bokmål*; neonorueguês (**nn**) — *nynorsk*.

Nota 5: Do alemão *Jiddisch*, a forma **iídixe** parece preferível a iídiche. ש e sch correspondem ao som /ʃ/ tradicionalmente representado por «x» em português (exemplos: xevá, para שווא/Schwa; xibolete, שִׁבֹּלֶת/Schibboleth).

1.1.4 *Línguas balto-eslavas*

Ramo da família indo-europeia, que inclui os sub-ramos eslavo e báltico. O sub-ramo eslavo inclui o diassistema servo-croata.

ISO 639	endónimo	pt	en	fr	IATE
be	беларуская мова	**bielorrusso** [DP, DM]	Belarusian	biélorusse	282193
bg	български	**búlgaro** [CRI]	Bulgarian	bulgare	282195
bs	bosanski jezik	**bósnio** [DP] bosníaco [DP] bosniano [DP] bosniense [DP]	Bosnian	bosniaque	338479
cs	čeština	**checo** [CRI] tcheco [PE, DP, AD, DM]	Czech	tchèque	282202
hr	hrvatski	**croata** [CRI] croácio [PE, AD, DM] croaciano [DM]	Croatian	croate	282214
lt	lietuvių kalba	**lituano** [CRI] lituânico [PE, DP, DE, AD, DM] lituânio [DM]	Lithuanian	lituanien	282231
lv	latviešu valoda	**letão** [CRI] lético [PE, DP, DM] leto [DM]	Latvian	letton	282232
mk	македонски јазик	**macedónio** [PE, DP, AD, DM]	Macedonian	macédonien	3531907
pl	polski	**polaco** [CRI] polonês [PE, DP, DE, AD, DM]	Polish	polonais	282246
ru	русский	**russo** [PE, DP, DE, AD, DM] russiano [DP, AD, DM]	Russian	russe	282251
sk	slovenčina slovenský jazyk	**eslovaco** [CRI]	Slovak	slovaque	1451616
sl	slovenščina slovenski jezik	**esloveno** [CRI] eslovénico [PE]	Slovenian	slovène	282257
sr	српски језик	**sérvio** [PE, DP, DE, AD, DM]	Serbian	serbe	282261
uk	українська	**ucraniano** [PE, DP, AD, DM] ucrânio [DM] ucraíno [DM] ruteno [PE, DP, AD, DM]	Ukrainian	ukrainien	282273
wen	serbsce serbski	**sorábio** [PE, DH] lusácio [PE, DH]	Sorbian languages	langues sorabes	223571

Nota 1: O **servo-croata** (**sh**) é uma língua eslava policêntrica/diassistema que inclui as seguintes variantes nacionais: bósnio (**bs**); montenegrino (**cnr**); croata (**hr**); sérvio (**sr**).

1.1.5 *Línguas albanesa, arménia e helénica*

Três ramos, atualmente cada um com uma única língua.

ISO 639	endónimo	pt	en	fr	IATE
el	ελληνικά	**grego** [CRI]	Greek	grec	—

hy	Հայերեն (hayeren)	**arménio** PE, DP, DE, AD, DM arménico PE, DE, AD, DM	Armenian	arménien	282216
sq	shqip	**albanês** PE, DP, DE, AD, DM	Albanian	albanais	282260

1.2 *Línguas uralo-altaicas*

Grupo de línguas, muitas vezes consideradas separadamente: línguas urálicas e línguas altaicas. A IATE inclui três línguas urálicas, todas elas línguas oficiais da União Europeia: o estónio, o finlandês e o húngaro. Das línguas altaicas, a IATE inclui várias línguas do ramo turco e o mongol. Inclui ainda o coreano e o japonês, cuja pertença a uma família altaica mais lata não é consensual.

ISO 639	endónimo	pt	en	fr	IATE
az	azərbaycan dili	**azerbaijano** PE, DP, DE, AD, DH, DA azeri PE, DP, DE, AD, DM, DH, DA azerbaidjano AD, DM, DH, DA azerbeijano VOLP azerbeidjano DA azerbaijanês PE, DP	Azerbaijani	azéri	282192
et	eesti keel	**estónio** CRI estoniano PE, DP, DE, AD, DM estónico PE, AD	Estonian	estonien	282205
fi	suomi	**finlandês** CRI finês PE, DP, DE, AD, DM fínico DP, AD, DM finense DM fino DM	Finnish	finnois	
hu	magyar	**húngaro** CRI magiar PE, DP, DE, AD, DM madgiar DM	Hungarian	hongrois	
ja	日本語 にほんご (nihongo)	**japonês** PE, DP, DE, AD, DM	Japanese	japonais	
kk	қазақ тілі	**cazaque** PE, DP, DE, AD, DM, DH, DA cazaquistanês PE, DP	Kazakh	kazakh	282223
ko	한국어 (hanguk-eo)	**coreano** PE, DP, DE, AD, DM	Korean	coréen	282226
mn	Монгол хэл	**mongol** PE, DP, DE, AD, DM mogol PE, DE, AD mongólico PE, DP, DE, AD, DM mogólico DE	Mongolian	mongol	282236
tr	Türkçe	**turco** PE, DP, DE, AD, DM, DH, DA	Turkish	turc	282272
uz	o‘zbekcha ўзбекча اۇزبېكچه	**usbeque** PE, DP, DH usbequistanês PE, DH uzbeque DP, DE, AD, DM uzbequistanês DP	Uzbek	ouzbek	282275

1.3 *Línguas esquimó-aleútes*

O grupo esquimó-aleúte corresponde a línguas faladas nas regiões árticas por populações autóctones, dispersas num vasto território, que vai desde o extremo leste da Sibéria até à ilha da Gronelândia. Essas línguas podem subdividir-se em aleúte e línguas iupiques, na vertente pacífica, e línguas inuítes, na vertente ártica e atlântica, as mais importantes em termos de número de falantes. A IATE inclui três línguas inuítes.

ISO 639	endónimo	pt	en	fr	IATE
ik	Iñupiaq Iñupiatun	**inupiaque**	Inupiaq	inupiaq	3581684
iu	ᐃᓄᒃᑎᑐᑦ (inuktitut)	**inuctitute** [ver nota 1] ***inuktitut*** [PE]	Inuktitut	inuktitut	3581794
kl	kalaallisut kalaallit oqaasii	**gronelandês** [PE, DP, DE] groenlandês [PE, DP, DE, AD, DM]	Kalaallisut Greenlandic	groenlandais	282224

Nota 1: Não havendo registo lexicográfico de nomes portugueses para a maioria das línguas inuítes, sugerem-se os aportuguesamentos **inupiaque** e **inuctitute**.

As **línguas inuítes**, subdividem-se, de oeste para leste, em:

- **ik** — inuíte alasquiano — **inupiaque**
- **iu** — inuíte canadiano:
  - inuíte oeste-canadiano (inuvialuctune)
  - inuíte leste-canadiano **(inuctitute)**
- **kl** — **gronelandês** (calalissute)

O termo **inuctitute** é frequentemente utilizado para representar o conjunto das variantes do sistema linguístico **inuíte canadiano**, não só as leste-canadianas como também as oeste-canadianas. A extrema dispersão das populações inuítes faz com que o inuctitute seja uma língua policêntrica por excelência.

2. LÍNGUAS AFRO-ASIÁTICAS

2.1 *Línguas afro-asiáticas*

Das línguas afro-asiáticas, a IATE inclui uma língua do ramo cuxita[(24)] (o somali), uma língua do ramo chádico (o haúça) e várias línguas do ramo semítico, entre as quais o maltês, língua oficial da UE.

ISO 639	endónimo	pt			IATE
am	አማርኛ (amarəñña)	**amárico** [PE, DP, AD, DA]	Amharic	amharique	282189
ar	اَلْعَرَبِيَّةُ (al-ʿarabiyyah)	**árabe** [PE, DP, DE, AD, DM, DA] arábico [PE, DP, DE, AD, DM, DA]	Arabic	arabe	282190
ha	هَرْشَن هَوْسَ (hausa)	**haúça** [PE, DP, AD, DH, DA] hauçá [DP, DM, DH, DA] haussá [AD]	Hausa	haoussa	3546094 282212
he	עברית (ivrit)	**hebraico** [PE, DP, DE, AD, DM, DA] hebreu [PE, DP, DE, DM, DA]	Hebrew	hébreu	282219
mt	Malti	**maltês** [CRI]	Maltese	maltais	
so	Soomaaliga af Soomaali	**somali** [PE, DE, AD, DM, DA] somáli [PE] somaliano [PE, AD, DM, DA]	Somali	somali	282259

3. LÍNGUAS SUL-SARIANAS

3.1 *Línguas nilo-sarianas*

ISO 639	endónimo	pt	en	fr	IATE
mas	ɔl Maa	**massai** [PE, DP, DE, AD, DM]	Masai	massaï	3581800

3.2 *Línguas nigero-congolesas*

Das línguas nigero-congolesas, a IATE inclui uma língua do ramo atlântico (o uolofe), uma língua do ramo mandinga (o bambara) e várias línguas do ramo banto.

ISO 639	endónimo	pt	en	fr	IATE
bam	ߓߡߊߣߊ߲ߞߊ߲ (bamanankan)	**bambara** [PE, DM, DH, DA]	Bambara	bambara	3581624
kmb	Kimbundu	**quimbundo** [PE, DP, DE, AD, DM] bundo [PE, AD] ambundo [DM] *kimbundu* [DP]	Kimbundu	kimbundu	3572162
ny	Chichewa, Chinyanja	**nianja** [PE, DE, DM, DH, DA; (ver nota 3)] cinianja [PE, DM, DH] cheua [DE, DA]	Chichewa Chewa Nyanja	chichewa chewa nyanja	1451603
rn	Ikirundi	**rundi** [DA] quirrundi	Rundi	rundi	3529011
rw	Ikinyarwanda	**ruanda** [DM, DA, DH; (ver nota 1)] **quiniaruanda** [PE] quiniarruanda	Kinyarwanda	rwanda	282252
ss	siSwati	**suáti** [(ver nota 4)] **suázi** [PE, DP, DE] suazi [DE, AD, DM, DH, DA] suazilandês [AD, DM, DH]	Swati	swati	918195
st	Sesotho	**soto** [DA; (ver notas 1 e 5)] **sesoto** [PE, DP, DE, DH] sessoto	Southern Sotho	sotho du Sud	282263
sw	Kiswahili	**suaíli** [PE, DP, DE, DM, DH, DA] suaíle [PE, DP, DE, AD, DM, DA] suaile [DH] quissuaíli [PE]	Swahili	swahili	282264
wo	Wolof	**uolofe** [PE, DP, DE, DA; (ver nota 6)] **jalofo** [PE, DP, DE, AD, DM, DH, DA] uólofe [DP, DM, DH] *uolof* [PE]	Wolof	wolof	282277
xh	isiXhosa	**xossa** [PE, DP, DA] *xhosa* [PE, DE] xosa [PE, DP, DA]	Xhosa	xhosa	282278
zu	isiZulu	**zulo** [PE, DP, DE, AD, DM, DH, DA; (ver nota 7)] **zulu** [DP, DE, AD, DM, DH, DA] *zulu* [PE] issizulo issizulu	Zulu	zoulou	918204

Nota 1: Nas línguas bantas, prefixos como *ki-*, *iki-*, *chi-*, *si-*, *isi-*, *se-*, etc. significam língua. Esses prefixos são, assim, «redundantes»:

- evitar: língua cinianja — preferível: língua **nianja**
- evitar: língua quirrundi — preferível: língua **rundi**
- evitar: língua quiniarruanda — preferível: língua **ruanda**
- evitar: língua sissuáti — preferível: língua **suáti** (ou suázi)
- evitar: língua sessoto — preferível: língua **soto**
- evitar: língua quissuaíli — preferível: língua **suaíli**
- evitar: língua issixossa — preferível: língua **xossa**
- evitar: língua issizulo — preferível: língua **zulo**

O **qui**mbundo é a exceção notável.

Nota 2: Convencionou-se na variante europeia do português transcrever o som /s/ intervocálico das línguas bantas com ss e não com o ç usado para palavras de línguas ágrafas ou que utilizam alfabetos exóticos. Exemplos: quissuaíli, xossa, Essuatíni, musseque, missanga, etc.

Nota 3: A mesma língua é designada **nianja** na margem moçambicana do lago Niassa e cheua na margem malauiana, onde é língua oficial. É de toda a lógica optar pela designação usada num país lusófono. O mesmo se poderá também dizer da opção pela designação **lago Niassa** em detrimento da designação lago Maláui.

Nota 4: O aportuguesamento **suáti** aproxima-se mais do endónimo *siSwati* — ficando em linha com o gentílico essuatiniano ou o topónimo Essuatíni, nova designação da Suazilândia.

Nota 5: O **soto** (**st**) é a variante meridional de um grupo de línguas aparentadas. O soto ocidental é conhecido como tsuana (**tn**) e o soto do norte como pedi (**nso**).

Nota 6: Jalofo é o antigo nome dado à língua e povo **uolofe**.

Nota 7: O termo **zulo** reflete a pronúncia [ˈzulu], mais próxima do original, enquanto a pronúncia de zulu [zuˈlu] se afasta do original. O mesmo se pode dizer de **quimbundo** ou de **umbundo**, línguas que nos estão historicamente mais próximas. Estas grafias são usadas por refletirem melhor em português a pronúncia banta do que as hipotéticas grafias quimbundu ou umbundu.

4. Línguas austricas

Macrofamília de línguas do Sudeste Asiático e do grande arco que vai de Madagáscar à Polinésia, passando pela Insulíndia.

4.1 *Línguas tai*

ISO 639	endónimo	pt	en	fr	IATE
lo	ລາວ (lao) ພາສາລາວ (phasa lao)	**laociano** PE, DP, DH, DA laosiano PE, DP, DE, AD, DH, DA lao PE, DA	Lao	lao	282230
th	ไทย (thai) ภาษาไทย (phasa thai)	**tailandês** PE, DP, DE, AD, DM, DH, DA tai DM, DA siamês DP, DE, DM, DH, DA siame PE, DP, DE	Thai	thaï	282268

4.2 *Línguas austro-asiáticas*

ISO 639	endónimo	pt	en	fr	IATE
km	ខ្មែរ (khmae) ភាសាខ្មែរ (phiăsaa khmae)	**cambojano** AD, DA; (ver nota 1) ***khmer*** PE, DP, AD, DH, DA cambodjano AD cambojiano DA cmere quemere	Central Khmer	khmer central	3588038
vi	Tiếng Việt	**vietnamita** PE, DP, DE, AD, DH, DA vietnamês PE, DM, DH vietnamiano PE, DP, DH vietnamense PE, DP, DM, DH	Vietnamese	vietnamien	282276

Nota 1: O **cambojano** (ou cambojano central) é a língua oficial do Camboja, sendo **cmere** ou **quemere** possíveis aportuguesamentos do endónimo ខ្មែរ (*khmae*).

4.3 *Línguas austranésias*

Família de línguas com possível origem na Formosa. Inclui o grande ramo malaio-polinésio com línguas faladas de Madagáscar, a ocidente, à ilha da Páscoa, a oriente, passando pela Insulíndia.

ISO 639	endónimo	pt	en	fr	IATE
fj	vosa Vakaviti	**fijiano** [PE, DP, DE, DM, DA] fidjiano [PE, DP, DM]	Fijian	fidjien	1451565
id	Bahasa Indonesia	**indonésio** [PE, DP, AD, DM] malaio-indonésio [DM] *bahasa* [DP] baassa	Indonesian	indonésien	282217
mg	fiteny malagasy	**malgaxe** [PE, DP, DE, DM] madagascarense [AD, DM]	Malagasy	malgache	282233
mi	te reo Māori	**maori** [PE, DP, DE, AD, DM]	Maori	maori	282234
ms	Bahasa Melayu بهاس ملايو	**malaio** [PE, DP, AD, DM]	Malay	malais	282238
sm	gagana fa'a Samoa	**samoano** [DP, AD, DM]	Samoan	samoan	282258
tet	lia-tetun	**tétum** [PE, DP, DE] teto [PE, DP, DE]	Tetum	tetum	344193
tl	Wikang Tagalog	**tagalo** [PE, DP, DE, AD, DM; (ver nota 2)] tagal [AD, DM]	Tagalog	tagalog	282269
to	faka Tonga	**tonganês** [DP, DE, DM] tonga[25] [PE, AD]	Tonga	tongan	
ty	reo tahiti	**taitiano** [PE, DP, DE, AD, DM]	Tahitian	tahitien	282284

Nota 1: भाषा (*bhāṣā*) é uma palavra do sânscrito, que significa idioma. Não é só o malaio/indonésio e outras línguas austranésias que usam esta palavra para se referir ao seu idioma, pois a influência do sânscrito faz-se notar noutras áreas linguísticas, das línguas indo-arianas (ex.: *hindī bhāṣā*) às sino-tibetanas (ex.: *myanma bhasa*), passando pelas dravídicas (ex.: *tamil bhāṣa*), cra-tai (ex.: *phasa thai*) e austro-asiáticas (ex.: *phiesa khmae*).[26] **Baassa** — possível aportuguesamento.

Nota 2: O **filipino** (**fil**) é a língua oficial das Filipinas, língua que tem por base o tagalo (**tl**).

5. Línguas ameríndias

Grupo de línguas com uma única língua na IATE.

5.1 *Línguas ameríndias*

ISO 639	endónimo	pt	en	fr	IATE
qu	Runa Simi Kichwa	**quíchua** [PE, DP, DE, AD, DM, DA] quéchua [PE, DP, DE, AD, DM, DA]	Quechua	quechua	282248

6. Línguas dene-caucasianas

Macrofamília contestada por muitos linguistas, que agrupa o basco, as línguas caucasianas, as línguas dravídicas e as línguas sino-tibetanas.

6.1 *Língua basca*

ISO 639	endónimo	pt	en	fr	IATE
eu	euskara	**basco** [PE, DP, AD, DM] vasco [DP, AD, DM] euscara [DP, DE] euscaro [AD, DM] êuscara [DE] êuscaro [PE, DE] eusquera [DP] vasconço [PE, DP, DE, AD, DM]	Basque	basque	146972

Nota 1: O **basco unificado** (*euskera batua*) é a norma criada como solução de compromisso entre as diferentes variantes do basco: biscainho, guipuscoano, navarro, labortano-navarro, suletino.

6.2 *Línguas caucasianas*

Conjunto de línguas faladas na região do Cáucaso.

ISO 639	endónimo	pt	en	fr	IATE
ab	аҧсуа бызшәа аҧсшәа	**abcásio** [PE, DP, DE] abcázio [DP]	Abkhazian	abkhaze	177971
ka	ქართული (kartuli ena)	**georgiano** [PE, DP, DE, AD, DM, DA] geórgico [AD]	Georgian	géorgien	282222

6.3 *Línguas dravídicas*

Faladas antigamente em todo o subcontinente indiano, estão hoje limitadas ao terço sul da península Hindustânica e ao norte do Seri Lanca.

ISO 639	endónimo	pt	en	fr	IATE
ta	தமிழ் (tamil)	**tâmil** [PE, DP, DE, AD, DM, DH, DA] tâmul [PE, DP, DE, AD, DM, DH, DA]	Tamil	tamoul	282265
te	తెలుగు (tĕlugu)	**télugo** [DP, AD, DM, DH; (ver nota 1)] **telugo** [PE, DH] telugu [DA] telinga [DP, AD, DM, DH] tuluva [DM, DH] túluva [AD] tulo [AD] túlu [DH]	Telugu	télougou	282266

Nota 1: **Télugo**, palavra esdrúxula, mais próxima da pronúncia original.

6.4 *Línguas sino-tibetanas*

Línguas tibeto-birmânicas e o chinês.

ISO 639	endónimo	pt	en	fr	IATE
bo	བོད་སྐད (bod skad)	**tibetano** [PE, DP, DE, AD, DM, DH, DA]	Tibetan	tibétain	282199

dz	རྫོང་ཁ (dzong-kha)	**butanês** PE, DP, AD, DM; (ver nota 1) **dzonga** DA butani DM, DH butâni PE, DH	Dzongkha	dzongkha	282203
my	ဗမာစာ (mranma ca)	**birmanês** PE, DP, DE, AD, DM, DH, DA	Burmese	birman	282240
zh	中文 (zhōngwén) 汉语 / 漢語 (hànyǔ)	**chinês** PE, DP, DE, AD, DM, DH, DA; (ver nota 2)	Chinese	chinois	887865

Nota 1: *Skad* em tibetano e *kha* em butanês significam língua. *Bod skad*, língua do Tibete, e *dzong-kha*, língua do palácio (*dzong*). Dzonga é também nome alternativo de tsua ou xítsua (Moçambique), código ISO tsc.

Nota 2: O chinês (**zh**) inclui um conjunto das línguas/variantes. O **mandarim** (**cmn**) é a variante mais falada, correspondendo-lhe o estatuto de língua oficial da República Popular da China. O **cantonês** (**yue**) é a variante mais falada em Macau e Hong Kong.

---

(1) Serviço de Publicações, *Código de Redação Interinstitucional*, «Anexo A5 — Lista dos Estados, territórios e moedas», http://publications.europa.eu/code/pt/pt-5000501.htm.
(2) Serviço de Publicações, *Código de Redação Interinstitucional* «Anexo A8 — Código das línguas (União Europeia)», http://publications.europa.eu/code/pt/pt-5000800.htm.
(3) Biblioteca do Congresso, *Standards: Codes for the Representation of Names of Languages*, https://www.loc.gov/standards/iso639-2/php/code_list.php.
(4) Candidatos: Albânia, Macedónia do Norte, Montenegro, Sérvia, Turquia. Potenciais candidatos: Bósnia-Herzegovina, Cossovo.
(5) Caso Montenegro opte por uma língua autónoma dentro do diassistema servo-croata.
(6) Língua desenvolvida pela Associação da Língua Auxiliar Internacional.
(7) Um primeiro levantamento já havia sido realizado no artigo de Santos, V. N. dos., «Línguas oficiais de países terceiros», «a folha», n.º 35 — primavera de 2011, https://ec.europa.eu/translation/portuguese/magazine/documents/folha35_pt.pdf.
(8) Porto Editora, *Dicionário Infopédia da Língua Portuguesa*, https://www.infopedia.pt/dicionarios/lingua-portuguesa.
(9) Priberam, *Dicionário Priberam da Língua Portuguesa*, https://dicionario.priberam.org/.
(10) Estraviz, I. A. (dir.), Durão, C. (dir.), *Dicionário Estraviz*, https://estraviz.org/.
(11) Lexikon, *Dicionário Aulete Digital*, http://www.aulete.com.br/.
(12) Michaelis, *Dicionário Brasileiro da Língua Portuguesa*, https://michaelis.uol.com.br/moderno-portugues/.
(13) Houaiss, A., Villar, M. de S., *Dicionário Houaiss da Língua Portuguesa*, Temas e Debates, Lisboa, 2003, 3 vol., ISBN 972-759-664-9.
(14) Ferreira, A. B. de H., *Novo Aurélio Século XXI: O Dicionário da Língua Portuguesa*, Nova Fronteira, 1999, 3.ª ed., ISBN 8520910106.
(15) Instituto de Linguística Teórica e Computacional, *Vocabulário Ortográfico do Português*, http://www.portaldalinguaportuguesa.org/?action=vop&page=info.
(16) Academia das Ciências de Lisboa, *Vocabulário Ortográfico da Língua Portuguesa*, https://volp-acl.pt/.
(17) Academia Brasileira de Letras, *Vocabulário Ortográfico da Língua Portuguesa*, https://www.academia.org.br/nossa-lingua/busca-no-vocabulario.
(18) Supersistema de língua construído a partir da análise estrutural de dous sistemas discretos e homogéneos que apresentam semelhanças parciais. Estraviz, I. A. (dir.), Durão, C. (dir.), *Dicionário Estraviz*, «Diassistema», https://estraviz.org/diassistema.
(19) Wikipedia, *List of ISO 639-1 codes*, https://en.wikipedia.org/wiki/List_of_ISO_639-1_codes.
Wikipedia, *List of ISO 639-2 codes*, https://en.wikipedia.org/wiki/List_of_ISO_639-2_codes.
(20) Em vez do código ISO **ht**, a IATE utiliza uma abreviatura de crioulo (**cr**).
(21) A não confundir com a variante escocesa do inglês, com o código **sco** (en: *Scots*; fr: *écossais*)
(22) Já dispôs de código próprio — **mo** —, entretanto abandonado.
(23) A não confundir com o ladino, a língua falada pelos sefarditas descendentes dos judeus da península Ibérica, de código ISO **lad**.
(24) De Cuxe, Núbia.
(25) Tonga ou guitonga é também o nome de língua banta de Moçambique, ISO: **toh**.
(26) Wikipedia, *Bhāṣā*, https://en.wikipedia.org/wiki/Bahasa.

Exoneração de responsabilidade: Os textos incluídos são da responsabilidade dos autores, não refletindo necessariamente a opinião da Redação nem das instituições europeias.
A Redação é responsável pela linha editorial d'«a folha», cabendo-lhe decidir sobre a oportunidade de publicação dos artigos propostos.

Redação: Paulo Correia (Comissão); Valdemar Ferreira (PE); Manuel Leal (Conselho da UE); Victor Macedo (CESE-CR); António Raúl Reis (Serviço das Publicações)
Grupo de apoio: Ana Luísa Faria (Conselho da UE); Hilário Leal Fontes (Comissão); Cristina Machado (Comissão); Susana Gonçalves (Comissão); Ana Lorenzo Garrido (Comissão); Joana Seixas (CESE-CR)
Paginação: Susana Gonçalves (Comissão)
Envio de correspondência: dgt-folha@ec.europa.eu

Edição impressa: oficinas gráficas do Serviço de Infraestruturas e Logística — Bruxelas (Comissão)
Edição eletrónica: sítio Web da Direção-Geral da Tradução da Comissão Europeia no portal da União Europeia — https://ec.europa.eu/translation/portuguese/magazine/pt_magazine_pt.htm

Centro de Informação Europeia Jacques Delors
DGAE/Ministério dos Negócios Estrangeiros

A coleção completa d'«a folha» está disponível no catálogo em linha da Biblioteca Jacques Delors
https://infoeuropa.eurocid.pt/pesquisar/wti=(a+folha)+AND+(wfmt=se+OR+wfmt=an)/catalogo=bibliografico

«a folha» ISSN 1830-7809

ISSN 1830-780-9
9771830780004